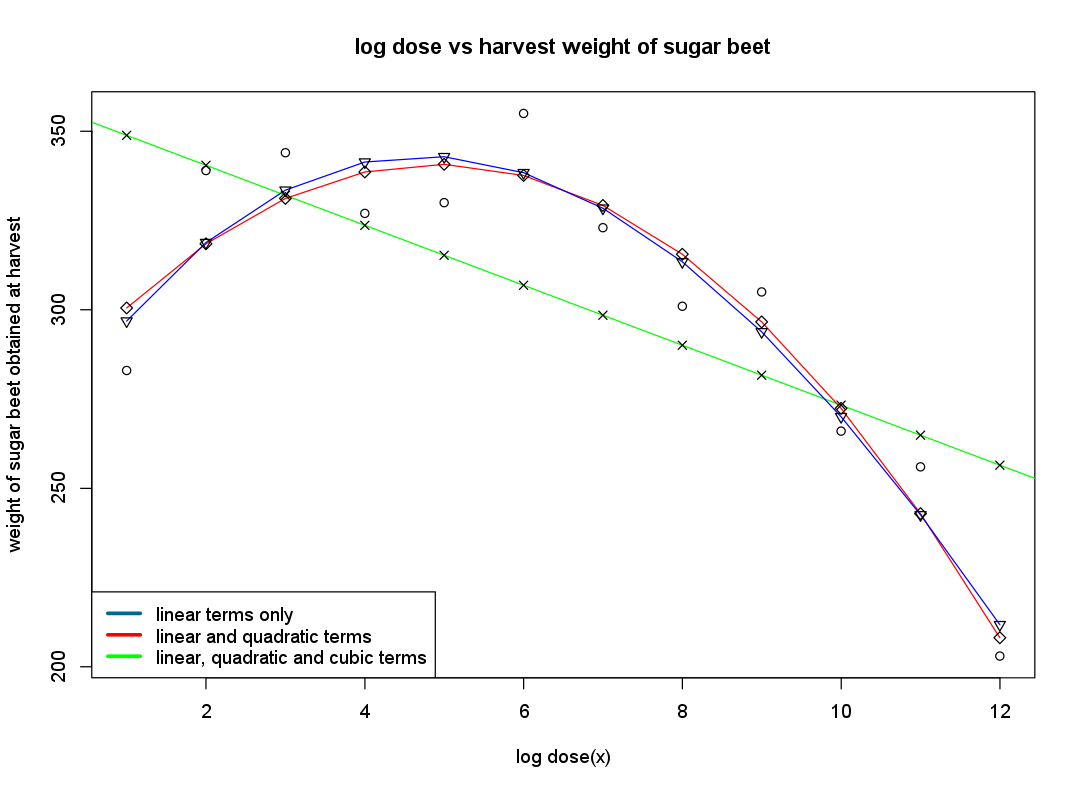
# Exercise 9



By visual inspection I can see that the quadratic and cubic models which include linear terms fit the model closest. Given that it is almost imperceivable the difference between the quadratic and the cubic model and the quadratic has fewer terms then I would prefer the quadratic model. My expectation is that the cubic model will fit better as it has mode degrees of freedom but will have to look at the adjusted R squared value to see if it is a better model.

Having looked at the output from R we can see that by analysing R-squared alone the best model is the cubic model (0.9092) but the quadratic was very close at (0.9062). However looking at adjusted R-squared we can see that the quadratic model would be a better choice with a value of 0.8854 vs 0.8752 for the cubic model.

|  |  |  |
| --- | --- | --- |
| Model | R-squared | Adjusted R-Squared |
| Linear only | 0.472 | 0.4192 |
| Linear + Quadratic | 0.9062 | 0.8854 |
| Linear + Quadratic + Cubic | 0.9092 | 0.8752 |

# Code for Exercise 9

x <- (1:12)  
y <- *c*(283, 339, 344, 327, 330, 355, 323, 301, 305, 266, 256, 203)  
  
beet\_data <- *data.frame*(x, y)  
  
linear\_only\_model <- *lm*(beet\_data$y ~ beet\_data$x)  
quadratic\_model <- *lm*(beet\_data$y ~ beet\_data$x + *I*(beet\_data$x^2))  
cubic\_model <- *lm*(beet\_data$y ~ beet\_data$x + *I*(beet\_data$x^2) + *I*(beet\_data$x^3))  
  
*plot*(beet\_data$x, beet\_data$y,  
 main="log dose vs harvest weight of sugar beet",  
 xlab="log dose(x)",  
 ylab="weight of sugar beet obtained at harvest")  
*abline*(linear\_only\_model, col="green")  
*lines*(beet\_data$x, quadratic\_model$fitted, col="red" )  
*lines*(beet\_data$x, cubic\_model$fitted, col="blue")  
*points*(beet\_data$x, linear\_only\_model$fitted, pch=4)  
*points*(beet\_data$x, quadratic\_model$fitted, pch=5)  
*points*(beet\_data$x, cubic\_model$fitted, pch=6)  
*legend*("bottomleft",*c*("linear terms only","linear and quadratic terms","linear, quadratic and cubic terms"),  
 col=*c*("deepskyblue4","red","green"), lwd=3)  
  
*summary*(linear\_only\_model)  
*summary*(quadratic\_model)  
*summary*(cubic\_model)

# Output of Code from Exercise 9

Call:

lm(formula = beet\_data$y ~ beet\_data$x)

Residuals:

Min 1Q Median 3Q Max

-65.859 -7.672 7.134 16.884 48.134

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 357.258 20.674 17.28 8.91e-09 \*\*\*

beet\_data$x -8.399 2.809 -2.99 0.0136 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 33.59 on 10 degrees of freedom

Multiple R-squared: 0.472, Adjusted R-squared: 0.4192

F-statistic: 8.939 on 1 and 10 DF, p-value: 0.01358

Call:

lm(formula = beet\_data$y ~ beet\_data$x + I(beet\_data$x^2))

Residuals:

Min 1Q Median 3Q Max

-17.517 -10.967 -5.681 12.888 20.514

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 277.2727 15.4221 17.979 2.32e-08 \*\*\*

beet\_data$x 25.8806 5.4545 4.745 0.001052 \*\*

I(beet\_data$x^2) -2.6369 0.4084 -6.456 0.000117 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.92 on 9 degrees of freedom

Multiple R-squared: 0.9062, Adjusted R-squared: 0.8854

F-statistic: 43.49 on 2 and 9 DF, p-value: 2.368e-05

Call:

lm(formula = beet\_data$y ~ beet\_data$x + I(beet\_data$x^2) + I(beet\_data$x^3))

Residuals:

Min 1Q Median 3Q Max

-14.383 -12.552 -4.756 11.733 20.180

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 267.16162 25.48032 10.485 5.95e-06 \*\*\*

beet\_data$x 33.69543 16.29346 2.068 0.0724 .

I(beet\_data$x^2) -4.08131 2.85380 -1.430 0.1906

I(beet\_data$x^3) 0.07407 0.14471 0.512 0.6226

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.57 on 8 degrees of freedom

Multiple R-squared: 0.9092, Adjusted R-squared: 0.8752

F-statistic: 26.7 on 3 and 8 DF, p-value: 0.0001611